Teilnehmen oder Boykottieren: ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx

Kühnel, Steffen M.; Jagodzinski, Wolfgang; Terwey, Michael

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Teilnehmen oder Boykottieren:
Ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx

von Steffen Kühnel, Wolfgang Jagodzinski und Michael Terwey[1]

In der empirischen Sozialforschung wird der Zusammenhang zwischen einer abhängigen und einer oder mehreren unabhängigen Variablen in der Regel mit Hilfe der (multiplen) linearen Regression untersucht. In vielen Fällen ist die abhängige Variable dichotom, hat also nur zwei mögliche Ausprägungen. Dann ist meistens weder die lineare Beziehung zwischen dieser und den erklärenden Variablen noch die Schätzung der Modellparameter nach der Methode der ungewichteten kleinsten Quadrate angemessen. Besser geeignet ist hier die Anwendung der logistischen Regression mit Maximum-Likelihood (ML) Schätzung der Modellparameter.

Obwohl die logistische Regression in der sozialwissenschaftlichen Methodenliteratur bereits seit längerer Zeit diskutiert wird (z.B. bei HANUSHEK und JACKSON, 1979; ARMINGER, 1983; ANDRESS, 1986), wird sie bei uns bislang erst selten eingesetzt. Eine Ursache hierfür mag in der, verglichen mit der linearen Regression, andersartigen und etwas komplizierteren Interpretation der Regressionskoeffizienten liegen. Daneben dürfte in der Vergangenheit aber auch die geringe Verfügbarkeit geeigneter Programme zur Schätzung logistischer Regressionsgleichungen die Anwendung begrenzt haben. Inzwischen ist allerdings die binäre logistische Regression mit ML-Schätzung in den drei großen und weit verbreiteten Programmpaketen BMDP (u.a. mit der Prozedur PLR), SAS (u.a. mit der Prozedur CATMOD) und SPSSx (mit PROBIT oder CNLR) standardmäßig verfügbar.

Wir wollen uns in diesem Beitrag etwas näher mit diesem Verfahren beschäftigen. Anhand eines einfachen Beispiels werden wir dazu das Grundkonzept der binären logistischen Regression, Tests der Effekte und die Interpretation der Resultate ausführlich diskutieren. Dabei ist unser Interesse primär auf die Darstellung der methodischen Aspekte gerichtet, weniger auf die inhaltliche Analyse, die wir entsprechend vereinfa-

---

1 Der vorliegende Beitrag basiert auf einer Zusammenfassung methodischer Überlegungen dreier stärker inhaltlich orientierter Arbeiten der Autoren (KÜHNEL, 1987; KÜHNEL und TERWEY, 1989; JAGODZINSKI und KÜHNEL, 1989).
Wolfgang JAGODZINSKI ist Professor an der Universität Gießen, die beiden anderen Autoren sind ZA-Mitarbeiter.

chen. Insbesondere werden wir auf Analogien, aber auch auf Unterschiede zur linearen Regression hinweisen. Da in der empirischen Sozialforschung z.Z. am häufigsten SPSSx eingesetzt wird, die Schätzung einer logistischen Regression mit SPSSx andererseits ein wenig umständlich ist, gehen wir im Anhang beispielhaft auf die Vorgehensweise bei einer Analyse mit SPSSx ein, um den Zugang etwas zu erleichtern.

1. Das Anwendungsbeispiel:
Affektive und instrumentelle Einstellung zur Volkszählung 1987 als Prädiktoren der Boykottabsicht

In unserem Anwendungsbeispiel soll die spontan geäußerte Absicht, die Volkszählung (VZ) 1987 zu boykottieren, erklärt werden. Prädiktoren sind die affektive Einstellung zur VZ und die Einschätzung der Notwendigkeit der Zählung. Die Daten wurden für die Begleituntersuchung zur Volkszählung 1987 erhoben (SCHEUCH u.a., 1988). Die Ergebnisse der Untersuchung werden demnächst in einer Publikationsreihe des Statistischen Bundesamtes veröffentlicht.

Im Rahmen der VZ-Begleituntersuchung ist in einer Befragung vor dem Beginn der VZ u.a. die **Boykottabsicht** bei der Zählung erfaßt worden. Die Frage sah vier Antwortmöglichkeiten vor: "Werde mich beteiligen", "Werde mich nicht beteiligen", "Kommt darauf an" und "Unentschieden, weiß noch nicht". Eingeschlossen in die folgende Analyse werden nur solche Personen, die die erste oder zweite Alternative genannt haben, sich[2] also dezidiert entweder für die Teilnahme oder für den Boykott ausgesprochen haben.

Erklärt werden soll diese abhängige Variable durch zwei Fragen nach den Einstellungen zur VZ. Bei der ersten Frage **VZ-Bewertung** wird auf einer siebenstufigen Skala mit den beiden Polen "Sehr ablehnend" (Antwortkode: 1) und "Sehr zustimmend" (Antwortkode: 7) die eher affektiv-emotionale Einstellung zur VZ erfaßt (vgl. SCHEUCH u.a., 1988). Demgegenüber soll die zweite Frage nach der **VZ-Notwendigkeit** stärker die kognitiv-instrumentelle Komponente erfassen. Dazu wurden die Befragten aufgefordert, sich für eines der drei Statements "Der Staat braucht genaue Statistiken; diese kann er nur durch eine Volkszählung erhalten.", "Der Staat braucht zwar genaue Statistiken; er kann diese aber auch ohne eine Volkszählung erhalten." oder "Der Staat braucht derartige Statistiken nicht." zu entscheiden. Die drei Wahlmöglichkeiten sind in einer dichotomen Variable mit den Kategorien "VZ ist notwendig" (1. Antwortalternative, Kode: 1)

2 Statistisch gesehen ist diese Reduktion auf nur 2 Kategorien der abhängigen Variable problematisch. Korrekt wäre eigentlich die Anwendung der Verallgemeinerung der binären logistischen Regression, die multinomiale logistische Regression.

und "VZ ist nicht notwendig" (2. und 3. Antwortalternative, Kode: 0) zusammengefaßt worden.

Tabelle 1: Die Häufigkeit der Boykottabsicht vor der VZ 1987

VZ-Notwendigkeit	Wert	VZ-Bewertung	Wert	Boykotteure N	(%)	Fallzahl
nicht notwendig	(0)	sehr ablehnend	(1)	118	(53.9)	219
nicht notwendig	(0)		(2)	36	(23.5)	153
nicht notwendig	(0)		(3)	10	(4.3)	235
nicht notwendig	(0)	neutral	(4)	2	(1.2)	168
nicht notwendig	(0)		(5)	1	(0.7)	138
nicht notwendig	(0)		(6)	1	(1.8)	60
nicht notwendig	(0)	sehr zustimmend	(7)	0	(0.0)	41
notwendig	(1)	sehr ablehnend	(1)	2	(10.0)	20
notwendig	(1)		(2)	1	(4.6)	22
notwendig	(1)		(3)	2	(2.0)	99
notwendig	(1)	neutral	(4)	0	(0.0)	126
notwendig	(1)		(5)	0	(0.0)	157
notwendig	(1)		(6)	0	(0.0)	209
notwendig	(1)	sehr zustimmend	(7)	0	(0.0)	414
				173	(8.4)	2061

(Quelle: VZ-Begleituntersuchung, 1. Panelwelle einschließlich Oversampling der VZ-Kritiker)

Befragt wurden im April und Mai 1987 1952 Personen einer repräsentativen Zufallsstichprobe. Das Sample wurde durch 604 zusätzliche Interviews mit Kritikern der VZ aus einer zweiten Zufallsstichprobe ergänzt (Oversampling von VZ-Kritikern). Aufgrund der Konzentration auf nur zwei der vier Antwortmöglichkeiten bei der Frage nach der Teilnahmeabsicht liegen von insgesamt 2061 der 2556 Befragten Antworten zu den drei oben genannten Variablen vor, die in Tabelle 1 zusammenfassend dargestellt sind.

2. Lineare und logistische Regression bei einer binären abhängigen Variable

Die Grundannahmen der logistischen Regression lassen sich u.E. am leichtesten in Gegenüberstellung zur üblichen linearen Regression verdeutlichen. Während bei einer deterministischen Funktion der Wert der abhängigen Variable exakt durch die Wertekom-

bination der unabhängigen Variablen bestimmt wird, gilt dies im nichtdeterministischen Regressionsmodell nur im Durchschnitt, d.h. für die Mittelwerte der abhängigen Variablen bei den verschiedenen Ausprägungskombinationen der unabhängigen Variablen. So werden im linearen Regressionsmodell die (bedingten) Mittelwerte in der Grundgesamtheit $\hat{y}$ einer abhängigen Variable Y als eine lineare Funktion der Werte der erklärenden Variablen $X_1, X_2, ..., X_k$ aufgefaßt:[3]

$$(1) \quad \hat{y} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + ... + \beta_k * X_k.$$

In unserem Anwendungsbeispiel ist die Boykottabsicht mit 1 und die Teilnahmeabsicht mit 0 kodiert. Der Mittelwert der Variable gibt hier also an, wie hoch der Anteil der potentiellen Boykotteure ist. Bezogen auf die Grundgesamtheit handelt es sich um die Wahrscheinlichkeit, daß eine Person boykottieren will. Aus Tabelle 1 ist zu entnehmen, daß dieser Anteil in der Stichprobe insgesamt 8,4% beträgt, der Mittelwert über alle Fälle also einen Wert von 0.084 hat.

Berechnet man die Anteile der potentiellen Boykotteure bzw. die Mittelwerte der abhängigen Variable jeweils für die Ausprägungskombinationen der unabhängigen Variablen, so schwanken die Anteile zwischen 53.9% bei Personen, die die VZ nicht für notwendig halten und affektiv sehr ablehnen, und 0% bei starken Befürwortern der VZ oder Personen, die eine neutrale oder positive Bewertung zur VZ haben und die Durchführung der Zählung für notwendig halten.

Abbildung 1 gibt die jeweiligen Mittelwerte bzw. Anteile der Boykotteure graphisch wieder. Wie man sieht, ist der Abfall der Boykotteuranteile mit steigender Sympathie zur VZ bei Personen, die die VZ nicht für notwendig halten, viel steiler als bei Personen, die die VZ für notwendig erachten. Im linear-additiven Regressionsmodell würde man daraus folgern, daß die VZ-Bewertung in der ersten Gruppe einen anderen Effekt als in der zweiten Gruppe hat, was durch die Spezifikation eines Interaktionseffekts zu berücksichtigen wäre.

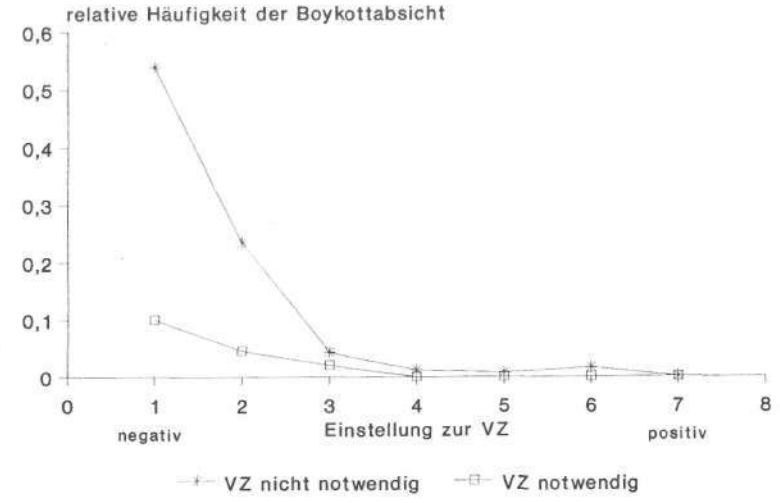
---

3 Der Einfachheit halber unterscheiden wir hier und im folgenden nicht explizit zwischen Grundgesamtheitswerten, Schätzern und Schätzungen. Statt $\hat{y}$ müßte hier eigentüch der Ausdruck:

$$\mu(Y|X_1 \cap X_2 \cap ... \cap X_k)$$

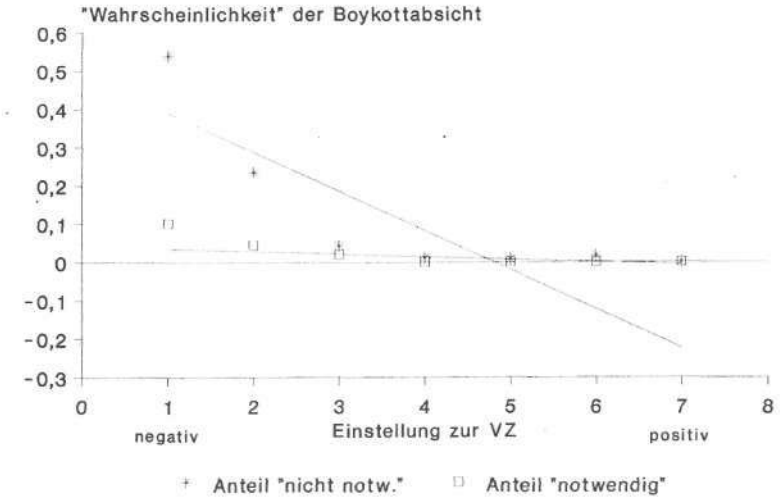
verwendet werden, da es sich bei der Gleichung (1) um bedingte Mittelwerte in der Grundgesamtheit handelt.

Abb. 1: Einfluß der Einstellung zur Volkszählung auf die Boykottabsicht



(Quelle: VZ-Begleituntersuchung, n=2061)

Abb. 2: Lineare Regression der Boykottabsicht mit Interaktionseffekt



(lineare Regression mit SPSSx 3.1)

Abbildung 2 zeigt das Ergebnis einer solchen linearen Regression. Die geschätzte Regressionsgleichung lautet:

$$(2) \qquad \hat{y} = 0.493 - 0.102 * X_1 - 0.450 * X_2 + 0.096 * X_1 * X_2,$$

wobei Y für die Boykottabsicht, $\hat{y}$ entsprechend für die geschätzten Grundgesamtheitsmittelwerte, X1 für die Bewertung der VZ und X2 für die Beurteilung der Notwendigkeit der VZ steht. Die zwei Geraden in Abbildung 2 ergeben sich durch die Aufsplittung der Regressionsgleichung nach der Beurteilung der Notwendigkeit der VZ:

$$(2a) \qquad \hat{y} = 0.493 - 0.102 * X_1 - 0.450 * 1 - 0.096 * X_1 * 1$$
$$= 0.043 - 0.198 * X_1 \qquad (\text{wenn } X_2 = 1)$$

bzw.

$$(2b) \qquad \hat{y} = 0.493 - 0.102 * X_1 \qquad (\text{wenn } X_2 = 0).$$

Die steilere Gerade in Abbildung 2 gibt also die Regression derjenigen an, welche die VZ nicht für notwendig halten, während die flachere Gerade die Regression derjenigen wiedergibt, welche die VZ für notwendig erachten.

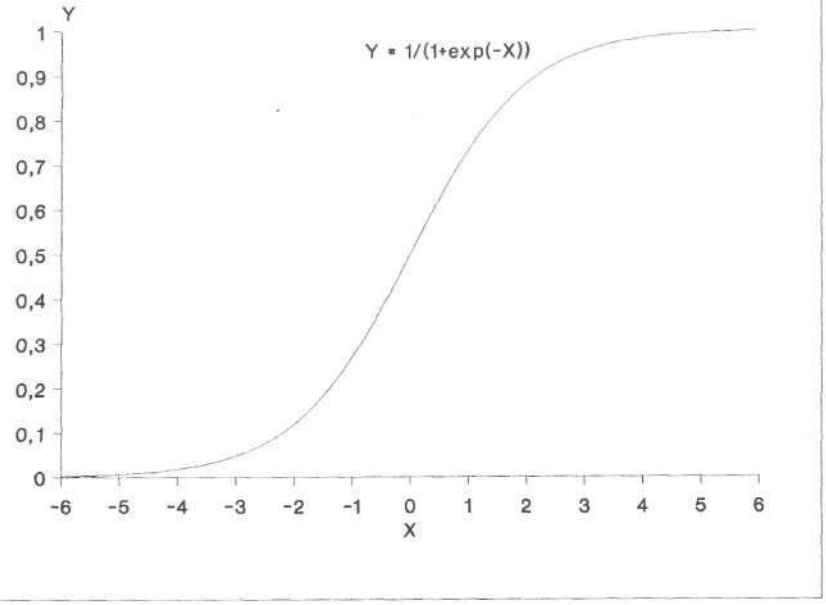
Abbildung 2 zeigt aber auch, daß die lineare Regression die empirischen Anteile nur recht unvollkommen approximiert. Sichtbar wird außerdem ein weiterer Nachteil der linearen Regression bei einer dichotomen abhängigen Variablen, daß nämlich "unmögliche" Werte auftreten können. Bei Personen, die die VZ nicht für notwendig halten ($X_2 = 0$), gibt die Regressionsgleichung bei einem Wert von X_1 von mindestens 5 negative Werte an. Da es Anteile bzw. Wahrscheinlichkeiten kleiner null nicht gibt, lassen sich solche Schätzwerte also schwerlich als bedingte Mittelwerte der 0/1-kodierten Variable interpretieren. Tatsächlich kann man schon Abbildung 1 entnehmen, daß die Verbindung der beobachteten Mittelwerte zu einer Kurve sowohl bei solchen Befragten, die die VZ für notwendig halten, als auch bei denen, die die VZ nicht für notwendig halten, einen deutlichen Knick aufweist. Die auf einer linearen Gleichung (Gerade) basierenden Schätzwerte dürften also kaum angemessen sein.

Günstiger erscheint hier die Verwendung einer nichtlinearen Regressionsfunktion. Diese Funktion sollte die Eigenschaft haben, bei beliebigen Wertekombinationen der unabhängigen Variablen nur Werte zwischen null und eins anzunehmen. Eine mathematische Funktion, die diese Bedingung erfüllt, ist die logistische Funktion, die durch die folgende Gleichung ausgedrückt werden kann:

(3) $Y = e^X / (1+e^X) = 1 / (1+e^{-X})$

e^X oder exp(X) ist die Exponentation zur Basis e, der Eulerschen Zahl (=2.718). Abbildung 3 zeigt die graphische Darstellung der S-förmigen Funktion. Bei steigenden Werten nähert sich die Kurve immer mehr dem Wert 1 an, bei fallenden Werten dem Wert 0.

Abb. 3: Logistische Funktion



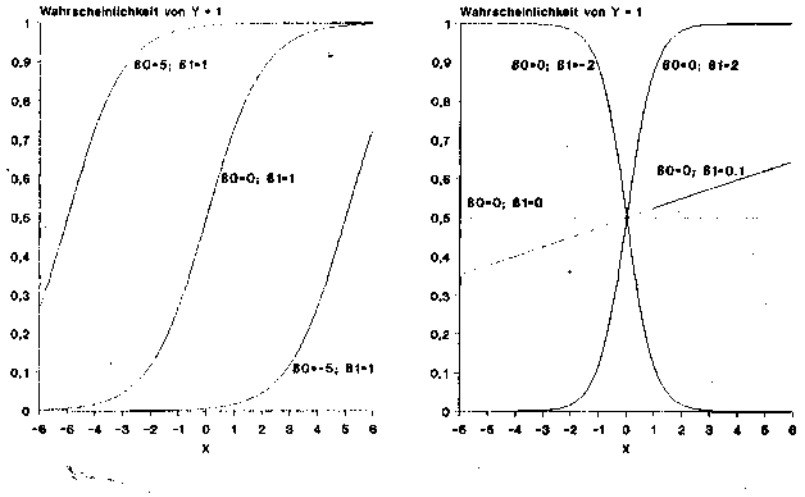
Im einfachsten logistischen Regressionsmodell werden analog zur bivariaten linearen Regression die Mittelwerte der dichotomen abhängigen Variable als logistische Funktion einer erklärenden Variable aufgefaßt. Flexibel wird die Gleichung durch die Einführung von Parametern. Die Regressionsfunktion hat dann die Form:

(4) $$\hat{y} = \exp(\beta_0 + \beta_1 \cdot X) / (1 + \exp(\beta_0 + \beta_1 \cdot X)) = 1 / (1 + \exp(-(\beta_0 + \beta_1 \cdot X)))$$

Abbildung 4 gibt einige Beispiele hierzu. Wie bei der linearen Regression ist $ß_0$ die Regressionskonstante und $ß_1$ das Regressionsgewicht der erklärenden Variable. Durch unterschiedliche Werte der Regressionskonstante wird die Kurve entlang der X-Achse verschoben, durch unterschiedliche Werte des Regressionsgewichts ändert sich die Steigung. Da die logistische Kurve in ihrem mittleren Bereich nahezu linear ist, lassen sich in diesem Bereich auch annähernd lineare Verläufe modellieren.

Abb. 4: Verschiedene logistische Regressionsfunktionen

$P(Y=1) = 1 / (1 + exp(-(ß_0 + ß_1 * X)))$



Bei mehr als einer unabhängigen Variable wird die logistische Funktion ganz analog zur linearen Funktion um weitere Regressionsgewichte ergänzt. Die allgemeine Form der Regressionsgleichung lautet dann:

$$(5) \quad \hat{y} = \frac{\exp(ß_0 + ß_1 * X_1 + ß_2 * X_2 + ... + ß_k * X_k)}{1 + \exp(ß_0 + ß_1 * X_1 + ß_2 * X_2 + ... + ß_k * X_k)}$$

$$= 1 / (1 + \exp(-(ß_0 + ß_1 * X_1 + ß_2 * X_2 + ... + ß_k * X_k)))$$

In der Gleichung (5) ist der Ausdruck, der exponiert wird, eine lineare Funktion der unabhängigen Variablen. Man kann die Gesamtgleichung daher auch in zwei Komponenten zerlegen, einen logistischen Teil:

$$(5a) \quad \hat{y} = e^{U} / (1+e^{U}) = 1 / (1+e^{-U})$$

und einen linearen Teil:

$$(5b) \quad U = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \ldots + \beta_k * X_k$$

Inhaltlich läßt sich das so deuten, daß die bei der Regression übliche lineare Funktion nicht direkt mit den Mittelwerten der abhängigen Variable verknüpft ist, sondern über eine sogenannte logistische Link-Funktion. Entsprechend wird die logistische Regression in der Statistik auch als ein Spezialfall des verallgemeinerten linearen Modells aufgefaßt (vgl. ARMINGER, 1983; ANDRESS, 1986).

Schätzt man nun mit der im nächsten Abschnitt besprochenen ML-Methode die Parameter der logistischen Regression der Boykottabsicht auf die affektive VZ-Einstellung und die Beurteilung der Notwendigkeit der Zählung, ergibt sich folgende Regressionsgleichung:

$$(6) \quad \hat{y} = \frac{\exp(1.581 - 1.436 * X_1 - 1.942 * X_2)}{1 + \exp(1.581 - 1.436 * X_1 - 1.942 * X_2)}$$

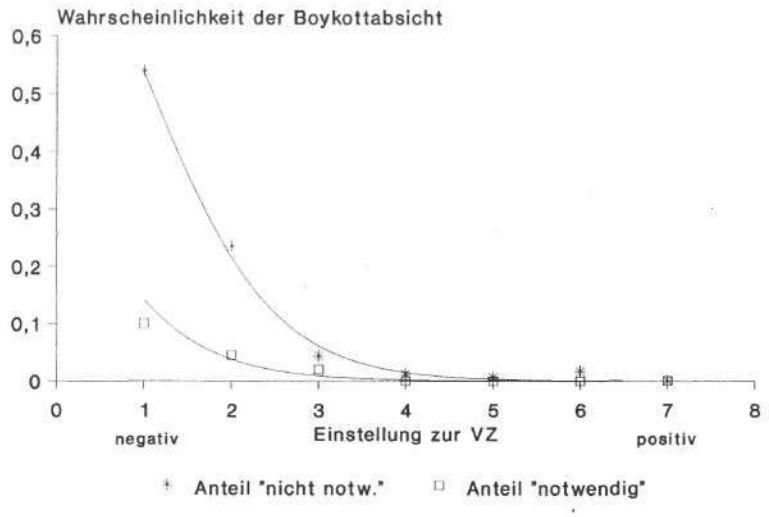
$$= 1 / 1 + \exp(-1.581 + 1.436 * X_1 + 1.942 * X_2)$$

Eine graphische Darstellung der logistischen Regressionsfunktion gibt Abbildung 5 wieder. Die obere Kurve zeigt die Regression bei Personen, welche die VZ nicht für notwendig halten, die untere Kurve die Regression bei Personen, welche die VZ für notwendig erachten. Verglichen mit der linearen Regression (Abbildung 2) ist die Übereinstimmung zwischen den beobachteten Mittelwerten und den nach der Regressionsfunktion erwarteten Mittelwerten deutlich besser, die Kurven weichen nur noch geringfügig von den empirischen Anteilen ab.[4] Die logistische Regression scheint daher - zumindest per Augenschein - für die vorgegebenen Daten deutlich geeigneter zu sein als die lineare Regression.

---

4 Bei einer optimalen Regressionsfunktion würden die Anteile der Boykotteure in allen durch die Wertekombinationen der unabhängigen Variablen aufgespannten Subpopulationen exakt vorausgesagt. Die Schätzung und Interpretation der Funktion erfolgt allerdings bei linearer und logistischer Regression auf der Ebene der einzelnen Fälle, im Unterschied zur Aggregatdatenanalyse etwa mit loglinearen Modellen.

Abb. 5: Logistische Regression der Boykott-Absicht



(CNLR mit SPSSx 3.1)

Bei der Gegenüberstellung der Regressionsgleichungen (2) und (6) fällt auf, daß die logistische Regression offenbar ohne einen Interaktionsterm auskommt, hier also das sparsamere Modell ist. Diese größere Sparsamkeit der logistischen Regression gegenüber der linearen findet man auch in anderen Anwendungen, etwa bei der Modellierung der Wahlabsicht bei Bundestagswahlen (vgl. JAGODZINSKI und KÜHNEL, 1989). Bevor wir im Zusammenhang mit der Interpretation von Ergebnissen der logistischen Regression auf diesen Punkt weiter eingehen, soll zunächst jedoch das Verfahren zur Schätzung der Parameter der Regressionsfunktion eingehender vorgestellt werden.

3. Maximum-Likelihood-Schätzung der Parameter der logistischen Regressionsfunktion

Bei der linearen Regression erfolgt die Schätzung der Modellparameter (Regressionskonstante und Regressionsgewichte) meist nach der Methode der ungewichteten kleinsten Quadrate (OLS-Methode, wobei OLS für "Ordinary Least Squares" steht). Die Parameter werden dabei so festgelegt, daß die Summe bzw. der Durchschnitt der quadrierten Differenzen aller beobachteten Werte der abhängigen Variable von den geschätzten Mittelwerten der Regressionsfunktion minimal ist. In die Schätzung der linearen Regressionsfunktion von Gleichung (2) gehen also alle 2061 Fälle der Stichprobe ein, ohne daß diese zuvor zu Gruppen gleichartiger Fälle zusammengefaßt werden.

Es läßt sich zeigen, daß die OLS-Methode unter bestimmten Bedingungen zu - im Sinne der statistischen Schätztheorie - guten Schätzungen führt (vgl. HANUSHEK und JACKSON, 1979). Diese Bedingungen sind nun allerdings bei einer dichotomen abhängigen Variable nicht gegeben. Bei der logistischen Regression wird daher üblicherweise ein anderes Schätzverfahren eingesetzt, die sogenannte Maximum-Likelihood (ML) Schätzung. Dabei werden die Parameter der Regressionsgleichung so festgelegt, daß die Wahrscheinlichkeit (bei kontinuierlichen Variablen die Wahrscheinlichkeitsdichte) maximal ist, die gegebenen Stichprobenwerte zu erhalten.

Wir haben oben erwähnt, daß bei einer 0/1-Kodierung der abhängigen Variable die bedingten Mittelwerte der abhängigen Variable als bedingte Wahrscheinlichkeiten für das Auftreten der mit 1 kodierten Ausprägung interpretiert werden können. Die Wahrscheinlichkeit eines Falles, der bei der abhängigen Variable den Wert 1 hat, ist also: [5]

$$(7a) \quad P(Y=1) = \hat{y} = 1 / (1 + \exp(-(\beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \ldots + \beta_k * X_k)))$$

Umgekehrt ist dann die Wahrscheinlichkeit eines Falles mit der Ausprägung 0 auf der abhängigen Variable eins minus dem Mittelwert:

---

5 Der Ausdruck P(Y = 1) steht hier und im folgenden als Abkürzung für eine bedingte Wahrscheinlichkeit bzw. die Schätzung einer Wahrscheinlichkeit, was korrekterweise durch die Formulierung:

$$P(Y=1 | X_1 \cap X_2 \cap \ldots \cap X_k)$$

ausgedrückt werden müßte (vgl. auch Fußnote 3).

$$(7b)\quad P(Y=0) = 1 - P(Y=1)$$
$$= 1 - \hat{y}$$
$$= 1 - 1/(1+\exp(-(\beta_0+\beta_1*X_1+\beta_2*X_2+...+\beta_k*X_k)))$$
$$= 1/(1+\exp(\beta_0+\beta_1*X_1+\beta_2*X_2+...+\beta_k*X_k))$$

Bei unabhängigen Realisierungen in der Stichprobe ist diese Gesamtwahrscheinlichkeit der Stichprobe das Produkt der Einzelwahrscheinlichkeiten. Die Gesamtwahrscheinlichkeit (Likelihood) der Stichprobe ist also:

$$(8)\quad P(Y_1) * P(Y_2) * ... * P(Y_n),$$

wobei $P(Y_i)$ die Wahrscheinlichkeit des aufgetretenen Y-Wertes des Falles i der insgesamt n Fälle der Stichprobe bezeichnet. In unserem Anwendungsbeispiel ist dies das Produkt der Wahrscheinlichkeiten der 2061 Fälle. Bei der Schätzung der Regressionskoeffizienten berechnet man allerdings nicht wirklich dieses Produkt, sondern stattdessen den zweifachen negativen Wert der Summe der logarithmierten Wahrscheinlichkeiten. Diese Funktion L ist zweimal das Negative der sogenannten Loglikelihood-Funktion:

$$(9)\quad L = -2 * \Sigma \ln(P(Y_i));\ i = 1,2,...,n$$

In Gleichung (9) wird für $P(Y_i)$ die jeweils zutreffende Funktion eingesetzt, also Gleichung (7a), wenn Yi = 1, und Gleichung (7b), wenn $Y_i = 0$. Die ß's werden dann so festgelegt, daß Gleichung (9) einen minimalen Wert annimmt. Es läßt sich zeigen, daß die Minimierung dieser Funktion gleichbedeutend mit der Maximierung aller Produkte von Wahrscheinlichkeiten in Gleichung (8) ist.

4. Interpretation der Ergebnisse einer logistischen Regression

Die Ergebnisse der Schätzung der Parameter der logistischen Regression unseres Anwendungsbeispiels sind in Tabelle 2 zusammengefaßt. Die Berechnungen wurden mit den Prozeduren PROBIT und CNLR in SPSSx 3.1 durchgeführt. Bei der folgenden Diskussion der Interpretation der Schätzung einer logistischen Regression werden wir wieder auf Analogien aber auch auf Unterschiede zur linearen Regression hinweisen.

Tabelle 2: Ergebnisse der logistischen Regression

Abhängige Variable: Boykottabsicht

-2 ln(Likelihood) des Nullmodells: 1188.32
-2 ln(Likelihood) des Endmodells: 644.97

Reduktion: $P^2 = 0.4574$ $chi^2 = 543.35$ $df = 2$

Erklärende Variablen	Regressions-koeffizient ß (t)	Effektkoeffizienten unstandard. exp(ß)	standard. exp(ß*s_x)	P^2-Zuwachs dP^2	(chi^2)
VZ-Bewertung	-1.436 (-12.7)	4.204^{-1}	17.857^{-1}	0.277	(328.9)
VZ-Notwendigkeit	-1.942 (- 4.1)	6.973^{-1}	2.639^{-1}	0.022	(25.6)
(Konstante)	1.581 (7.5)				

(Daten: VZ-Begleituntersuchung; n = 2061
Programm: SPSSx 3.1: PROBIT u. CNLR)

4.1. Schätzung und Interpretation von bedingten Anteilen in der Grundgesamtheit

Wie bei der linearen Regression lassen sich auch bei der logistischen Regression über die Regressionsparameter ß die Mittelwerte der abhängigen Variable für jede Ausprägungskombination der unabhängigen Variablen berechnen. Der geschätzte Anteil der Boykotteure bei Personen, welche die VZ für notwendig halten ($X_2 = 1$) und die eine neutrale Bewertung gegenüber der VZ einnehmen ($X_1 = 4$), berechnet sich z.B. durch Einsetzen der Werte der erklärenden Variablen in die geschätzte Regressionsfunktion (6) als:

$$(10a) \quad P(Y=1) = 1/(1 + \exp(-1.581 + 1.436 * 4 + 1.942 * 1)) = 0.0022267 \approx 0.2\%.$$

Für Personen, die eine sehr ablehnende Bewertung zur VZ einnehmen ($X_1 = 1$) und die Zählung nicht für notwendig erachten ($X_2 = 0$), beträgt der geschätzte Anteil der Boykotteure dagegen:

$$(10b) \quad P(Y=1) = 1/(1 + \exp(-1.581 + 1.436 * 1 + 1.942 * 0)) = 0.5361866 \approx 53.6\%.$$

In der linearen Regression werden die nach der Regressionsgleichung geschätzten Mittelwerte in der Regel als Vorhersagewerte interpretiert, mit denen man Werte der abhängigen Variable prognostizieren kann. Bei einer dichotomen abhängigen Variablen macht diese Interpretation wenig Sinn, da nur die Werte 0 und 1 auftreten können, ein Prognosewert von z.B. 0.536 also niemals realisiert werden kann. Dies bedeutet aber nicht, daß man die Mittelwerte nicht auf der Individualebene interpretieren kann und sie nur als Anteile der Boykotteure bei verschiedenen Merkmalskombinationen der erklärenden Variablen auffassen darf. Die Werte auf der Regressionsfunktion lassen sich vielmehr auch als individuelle Wahrscheinlichkeiten interpretieren, im Beispiel also als Wahrscheinlichkeit, daß eine Person mit dieser Ausprägungskombination die VZ boykottieren will.

4.2. Unstandardisierte Effektkoeffizienten

Bei Regressionsanalysen in der Sozialforschung geht es oft weniger um die Prognose (Schätzung von Mittelwerten) als um Aussagen über die Stärke des Einflusses der erklärenden Variablen auf die abhängige Variable. Hierbei ist zu beachten, daß die Effekte erklärender Variablen bei der logistischen Regression anders als bei der linearen Regression zu interpretieren sind. Zwar geben auch hier die Regressionskoeffizienten die Richtung des direkten Effekts einer erklärenden Variable an, und es gilt generell, daß die Änderung um eine Einheit einer erklärenden Variable um so stärker ist, je größer der absolute Wert des Regressionsgewichtes ist. Aber die Werte lassen sich nicht einfach in erwartete Änderungen bei der abhängigen Variable übersetzen, hier also in Wahrscheinlichkeits- bzw. Anteilsveränderungen.

Wenn etwa ein Befragter seine Bewertung der VZ um eine Einheit (+1 Einheit) in Richtung auf eine größere Befürwortung der VZ ändert, dann bewirkt diese Einstellungsänderung bei einer Person, die die VZ für notwendig hält und zunächst eine neutrale Bewertung gegenüber der VZ hat, daß die Boykottwahrscheinlichkeit von 0.2% (nach Gleichung 10a) um nur 0.15% auf 0.05% sinkt, da sich nun die Boykottwahrscheinlichkeit nach

$$(10c) \quad P(Y=1) = 1/(1+\exp(-(1.581-1.436*5-1.942*1))) = 0.0005305 \approx 0.05\%$$

berechnet. Ändert dagegen ein Befragter, der die VZ nicht für notwendig hält und die Zählung zunächst sehr stark ablehnt, seine VZ-Bewertung ebenfalls um +1 Einheit in

eine nicht mehr ganz so starke Ablehnung der Zählung, so sinkt seine Boykottwahrscheinlichkeit um immerhin 32% von 53.6% (nach Gleichung 10b) auf 21.6%:

(10d) $P(Y = 1) = 1 / (1 + \exp(-(1.581 - 1.436 * 2 - 1.942 * 0)))$
$= 0.2156836 \approx 21.6\%.$

Die Veränderung der abhängigen Variable bei einer Werteverschiebung auf einer unabhängigen Variable ist also von der jeweiligen Ausgangslage der abhängigen Variable abhängig. Dies liegt an der unterschiedlichen Steilheit der logistischen Regressionsfunktion. Bei Ausgangswahrscheinlichkeiten in der Nähe von 0 und 1 verläuft die Regressionskurve viel flacher als bei mittleren Wahrscheinlichkeiten. Da diese Ausgangslage nun auch von den Werten der übrigen erklärenden Variablen mitbestimmt ist, wird bei der logistischen Regression gewissermaßen implizit ein Interaktionseffekt berücksichtigt (LONG, 1987). Dies erklärt auch, warum bei einer logistischen Regressionen oft auf einen expliziten Interaktionsterm verzichten werden kann, selbst wenn ein solcher zusätzlicher Parameter bei einer linearen Regressionen notwendig wäre.

Unabhängig von der jeweiligen Ausgangslage ist jedoch bei einer Änderung einer unabhängigen Variable um +1 Einheit die Änderung des Verhältnisses der Wahrscheinlichkeiten bzw. der Anteile der beiden Ausprägungen der abhängigen Variable. Bei der Wertekombination der erklärenden Variablen von Gleichung (10a) beträgt das Verhältnis der Wahrscheinlichkeit von Boykottabsicht zu Teilnahmeabsicht 1 zu 448:

(11a) $P(Y = 1) / P(Y = 0) = 0.0022267 / (1 - 0.0022267) \approx 1 / 448$

Bei der nach Gleichung (10c) berechneten Wahrscheinlichkeit beträgt das Verhältnis von Boykott- zu Teilnahmewahrscheinlichkeit 1 zu 1848:

(11b) $P(Y = 1) / P(Y = 0) = 0.0005305 / (1 - 0.0005305) \approx 1 / 1848$

Das Wahrscheinlichkeitsverhältnis sinkt bei einer Änderung der VZ-Bewertung um +1 Einheit also um den Faktor 0.238:

(11c) $1 / 1848 \approx 1 / 448 * 0.238$

Um den gleichen Faktor ändert sich auch das Wahrscheinlichkeitsverhältnis bei der Änderung von (10b) nach (l0d). Nach Gleichung (10b) ergibt sich ein Verhältnis von 1.156 zu 1:

(11d) $P(Y = 1) / P(Y = 0) = 0.5361866 / (1 - 0.5361866) \approx 1.156 / 1$

Nach Gleichung (lOd) ist das Verhältnis 1 zu 3.636:

(11e) $P(Y=1) / P(Y=0) = 0.2156836 / (1-0.2156836) \approx 1 / 3.636$

Die Änderung um + 1 Einheit bewirkt wieder eine Veränderung des Wahrscheinlichkeitsverhältnisse um den Faktor 0.238:

(11f) 1 / 3.636 ≈ 1.156 / 1 • 0.238

Der Faktor 0.238, um den sich das Wahrscheinlichkeitsverhältnis ändert, ist gerade die Exponentation des Regressionskoeffizienten der VZ-Bewertung ($e^{-1.436} = 0.238$). Tatsächlich läßt sich zeigen, daß in der logistischen Regression generell gilt:

> *Bei der binären logistischen Regression bewirkt die Erhöhung einer unabhängigen Variable Xk um +1 Einheit, daß sich das Verhältnis der Wahrscheinlichkeiten der Ausprägungen der abhängigen Variable [P(Y=1) /P(Y=0)] um genau exp(ßk) ändert.*

Aufgrund dieser Beziehung hat LONG (1987) vorgeschlagen, bei einer logistischen Regression die von ihm als (unstandardisierten) Effektkoeffizienten $E(X_i)$ bezeichnete Exponentation der Regressionskoeffizienten zu betrachten:

(12) Effektkoeffizient: $E(X_i) = \exp(\beta_i)$

Tabelle 2 ist zu entnehmen, daß dieser Koeffizient bei der VZ-Bewertung den schon erwähnten Wert 0.238 bzw. 1 / 4.204 annimmt, während er bei der VZ-Notwendigkeit 0.143 (= 1 / 6.973) beträgt.[6] Bei der Interpretation dieser Koeffizienten ist zu beachten, daß die Effektkoeffizienten im Unterschied zu den Effekten bei der linearen Regression multiplikativ und nicht additiv wirken. Dies bedeutet, daß mehrfache Änderungen nicht einfach addiert werden können, sondern multipliziert werden müssen. Steigt etwa in unserem Beispiel die Einstellung zur VZ um +1 Einheit, ändert sich das Wahrscheinlichkeitsverhältnis um den Faktor 0.238. Steigt der Koeffizient um eine weitere Einheit, ändert sich das Verhältnis wiederum um den Faktor 0.238. Die Gesamtänderung, also die Änderung um + 2 Einheiten, beträgt nun nicht 2 ∗ 0.238 sondern 0.238 ∗ 0.238 = 0.238^2 = 0.057.

6 Bei Effektkoeffizienten kleiner 1 ist es günstiger, den Kehrwert des Koeffizienten zu betrachten. In Tabelle 2 sind entsprechend die Kehrwerte eingetragen.

Die Multiplikativität des Effektes ist natürlich auch bei einer negativen Veränderung zu berücksichtigen. Sinkt etwa die Einstellung zur VZ um -1 Einheit, so ändert sich das Verhältnis der Wahrscheinlichkeiten von Boykottabsicht zu Teilnahmeabsicht nicht um den Faktor -0.238 sondern um den Faktor 1 / 0.238 = 4.204.

Die Multiplikativität des Effektkoeffizienten wirkt sich auch auf die Interpretation der Größen der Koeffizienten aus. Ein Effektkoeffizient von 1 bedeutet, daß **kein** Effekt vorliegt, da sich das Wahrscheinlichkeitsverhältnis nicht ändert. Dieser Effektkoeffizient korrespondiert mit dem Regressionsgewicht 0 ($e^0 = 1$). Effektkoeffizienten größer 1 bedeuten, daß sich das Wahrscheinlichkeitsverhältnis der abhängigen Variable zugunsten der Ausprägung mit dem Wert 1 ändern, die korrespondierenden ß's sind dann positiv. Effektkoeffizienten kleiner 1 bedeuten dagegen, daß sich das Wahrscheinlichkeitsverhältnis zu Ungunsten der Ausprägung mit dem Wert 1 ändert. Der Effekt ist dann negativ.

Um positive mit negativen Effekten vergleichen zu können, ist es sinnvoll, bei Effektkoeffizienten kleiner eins den Kehrwert anzugeben. Dadurch ist es möglich, die Stärke der Effekte zu vergleichen. Beträgt etwa ein Effektkoeffizient einer unabhängigen Variable 3.6 und der einer zweiten Variable 0.25 ($= 1/4 = 4^{-1}$), so bedeutet dies, daß sich bei einer Änderung um +1 Einheit bei der ersten Variable das Wahrscheinlichkeitsverhältnis um den Faktor 3.6 zu gunsten der mit dem Wert 1 kodierten Ausprägung der abhängigen Variable ändert. Bei der zweiten Variable ändert sich dagegen das Wahrscheinlichkeitsverhältnis bei einer Änderung um +1 Einheit um den Faktor 4 zu Ungunsten der mit dem Wert 1 kodierten Ausprägung der abhängigen Variable. Die zweite Variable hat also im Vergleich zur ersten Variable einen größeren Effekt, der allerdings in die andere Richtung geht.

In unserem Anwendungsbeispiel sind die Effektkoeffizienten beider Prädiktoren kleiner eins, die Effekte also, was auch die Regressionskoeffizienten zeigen, negativ. Bei einer Änderung um +1 Einheit ändert sich das Wahrscheinlichkeitsverhältnis also zu Ungunsten der Boykottabsicht. Die VZ-Notwendigkeit hat hierbei einen stärkeren unstandardisierten Effekt als die VZ-Bewertung. So sinkt das Wahrscheinlichkeitsverhältnis bei einer Änderung der VZ-Bewertung um +1 Einheit um mehr als das vierfache (1 / 4.204), während die Änderung der VZ-Notwendigkeit um +1 Einheit ein Sinken um fast das siebenfache (1 / 6.973) bewirkt.

4.3. Standardisierte Effektkoeffizienten

Bei einem Vergleich der Effektstärke unter den erklärenden Variablen sollte auch die Skalierung der Variablen berücksichtigt werden. So ist in unserem Anwendungsbeispiel die Gesamtvariation der VZ-Bewertung viel größer als die Variation der VZ-Notwendigkeit, da letztere nur 2 Ausprägungen hat, die VZ-Bewertung jedoch 7. Man kann die unterschiedliche Stichprobenvariation berücksichtigen, indem man nicht die Veränderung um +1 Einheit als Maß nimmt, sondern die Veränderung um +1 Standardabweichung der jeweiligen erklärenden Variable. Man erhält so den standardisierten Effektkoeffizienten ${}_sE(X_k)$ von LONG (1987):

(13) standardisierter Effektkoeffizient:

$$ {}_sE(X_k) = \exp(\beta_k * s_k) = E(X_k) * \exp(s_k) $$

In Gleichung (13) bezeichnet sk die Standardabweichung der unabhängigen Variable X_k, die z.B. mit der Prozedur DESCRIPTIVES in SPSSx berechnet werden kann. Nach Tabelle 2 ergibt sich für die VZ-Bewertung ein Wert von 1 / 17.857 (= 0.056) und für die VZ-Notwendigkeit ein Wert von 1 / 2.639 (= 0.379). Bei einer Änderung um eine Standardabweichung sinkt also das Verhältnis von Boykott- zu Teilnahmewahrscheinlichkeit bei der VZ-Bewertung um fast das 18-fache, während es bei der VZ-Notwendigkeit um weniger als das 3-fache sinkt. Bezogen auf die standardisierten Effektkoeffizienten hat nun die VZ-Bewertung einen sehr viel größeren Effekt auf das Wahrscheinlichkeitsverhältnis als die VZ-Notwendigkeit.

4.4. Die Messung der Gesamterklärungskraft durch Pseudo-R^2

Neben der Größe der direkten Effekte interessiert die Gesamterklärungskraft eines Regressionsmodells. Der in der linearen Regression verwendete Determinationskoeffizient R^2 ist für das logistische Regressionsmodell ungeeignet, da im Unterschied zur linearen Regression die Residuen und die Werte auf der Regressionskurve nicht notwendigerweise unkorreliert sind. In unserem Anwendungsbeispiel korrelieren die nach der Regressionsfunktion geschätzten Boykottwahrscheinlichkeiten und die Residuen zwar nur mit dem Wert 0.0076, in anderen Fällen kann die Korrelation jedoch höher sein.

Anstelle des Determinationskoeffizenten kann man bei der logistischen Regression jedoch ein alternatives Zusammenhangsmaß verwenden, daß in mancher Hinsicht ein Analogon zum Determinationskoeffizienten ist. Bei der OLS-Schätzung der linearen Re-

gression wird der Durchschnitt der quadrierten Residuen (s_e^2) minimiert. Die Varianz der abhängigen Variablen (s_y^2) ist der Durchschnitt der quadrierten Residuen vom Mittelwert, entspricht also einer linearen Regression mit einer Regressionskonstante, aber ohne Prädiktoren. Der Determinationskoeffizient R^2 (= $1 - s_e^2/s_y^2$) gibt nun den Anteil an, um den sich der Durchschnittswert der quadrierten Residuen vermindert, wenn zusätzlich zur Regressionskonstante alle Prädiktoren in die lineare Regressionsgleichung aufgenommen werden. Ein ganz analoges Maß läßt sich für die ML-Schätzung der logistischen Regression konstruieren, wenn man statt dem Durchschnitt der quadrierten Abweichungen das zweifache Negative der logarithmierten Likelihoodfunktion aus Gleichung (9) verwendet:

(14) $P^2 = 1 - L_a / L_k$

In Gleichung (14) ist L_a der Wert der Minimierungsfunktion (9) bei einem Modell mit allen Prädiktoren, während L_k der Wert der Minimierungsfunktion von einem Modell ist, daß nur die Regressionskonstante $ß_0$, aber keine Prädiktoren enthält.

Das von uns als P^2 symbolisierte Zusammenhangsmaß wird auch Pseudo-R^2 genannt (vgl. HENSHER und JOHNSON, 1981).[7] Der Wertebereich von P^2 läuft von 0 bis 1, wobei der Wert 0 auftritt, wenn sich durch die erklärenden Variablen der Minimierungswert nicht verbessert. Dann sind auch die Schätzwerte aller Regressionsgewichte der erklärenden Variablen null und es besteht kein (monotoner) Zusammenhang zwischen den unabhängigen Variablen einerseits und der abhängigen Variable andererseits. Der Maximalwert 1 würde erreicht werden, wenn die Minimierungsfunktion (9) den kleinstmöglichen Wert, nämlich null, annimmt. Die geschätzten Wahrscheinlichkeiten der beiden Ausprägungen der abhängigen Variable würden dann nur die Werte 0 und 1 aufweisen, und innerhalb jeder Ausprägungskombination der erklärenden Variablen gäbe es keine Streuung.

Für die Verwendung von P^2 und gegen die Verwendung von R^2 sprechen auch inferenzstatistische Erwägungen. Während nämlich bei P^2 statistische Tests zulässig sind (vgl. den nächsten Abschnitt) ist der F-Test für R^2 bei einer dichotomen abhängigen Variable unzulässig.

---

7 In der Literatur werden allerdings auch noch andere Zusammenhangsmaße als Pseudo-R^2 bezeichnet (vgl. ALDRICH und NELSON, 1984).

Neben der Gesamterklärungskraft läßt sich die Differenz von P^2-Werten auch zur Beurteilung der Höhe von Netto-Effekten der erklärenden Variablen heranziehen. Dazu berechnet man zunächst P^2 für eine logistische Regression ohne eine spezielle, erklärende Variable in der Regressionsgleichung und anschließend P^2 für eine zweite Regression, in der zusätzlich diese Variable in die Gleichung aufgenommen wird. Je größer der Zuwachs von P^2 ist, um so stärker ist der Netto-Effekt dieser erklärenden Variable. Auch dieser variablenspezifische Zuwachs von P^2 läßt sich statistisch testen.

In Tabelle 2 sind neben den Regressionskoeffizienten und den Effektkoeffizienten nach LONG auch P^2 und die P^2-Zuwächse der beiden Prädiktoren aufgeführt. Die Gesamterklärungskraft beträgt knapp 46% (P^2 = 0.4574), was bei nur zwei erklärenden Variablen in der Regression ein sehr hoher Wert ist. Die Boykottabsicht läßt sich also bereits recht gut durch die beiden Prädiktoren vorhersagen. Bei der Interpretation ist allerdings zu berücksichtigen, daß es sich hierbei um eine spontane Antwort auf die Frage nach dem beabsichtigten Verhalten handelt und nicht um das tatsächliche spätere Verhalten. Vergleicht man neben dem absoluten Wert von P^2 den relativen Zuwachs bei den beiden erklärenden Variablen, so ist der Netto-Effekt der VZ-Bewertung stärker als der der VZ-Notwendigkeit.

5. Statistische Tests

Ähnlich wie bei der linearen Regression lassen sich auch die Regressionskoeffizienten der logistischen Regression inferenzstatistisch absichern. So können bei der ML-Schätzung die Standardfehler (geschätzte Standardabweichungen) der Parameter der logistischen Regressionsgleichung berechnet werden. Wenn bi der Schätzwert des Regressionskoeffizienten ßi und s(ßi) der Standardfehler von ßi ist, dann gilt, daß die Größe

$$(15) \quad (b_i - \beta_i) / s(\beta_i)$$

asymptotisch standardnormalverteilt ist (vgl. auch LONG, 1987). Um die Hypothese Ho: ßi = 0 zu testen, dividiert man den geschätzten Regressionskoeffizienten durch seinen Standardfehler. In Tabelle 2 sind die resultierenden Werte der üblichen Konvention folgend als t-Werte bezeichnet worden. Die Hypothese Ho wird im zweiseitigen Test mit der Irrtumswahrscheinlichkeit *a* verworfen, wenn die Testgröße t größer ist als $z_{1-\alpha/2}$ bzw. kleiner als $z_{\alpha/2}$, wobei z das entsprechende Quantil der Standardnormalverteilung bezeichnet. Nach der üblichen Faustregel ist ein Regressionskoeffizient dann signifikant von null verschieden, wenn t dem Betrag nach größer als 2 ist. Wie Tabelle 2 zeigt, sind danach alle Regressionskoeffizienten signifikant.

Da nicht nur die Standardfehler sondern auch die Korrelationen der Schätzfunktionen geschätzt werden können, läßt sich auch testen, ob mindestens einer von mehreren Regressionskoeffizienten signifikant ist. In der Regel verwendet man als Prüfgröße für den Einfluß mehrerer Prädiktoren jedoch nicht eine Verallgemeinerung von Gleichung (15) sondern die Differenz der Minimierungsfunktionen (9) zweier Regressionsmodelle. Wenn nämlich die Schätzung zweier logistischer Regressionsfunktionen auf denselben Stichprobenfällen beruht und eine der beiden Regressionsfunktionen eine Teilmenge der Prädiktoren der anderen Funktion ist, dann ist die Differenz der Minimierungsfunktionen (9) bei ineinander geschachtelten Modellen asymptotisch chiquadratverteilt. Dieser sogenannte Likelihood-Ratio-Test[8] prüft also, ob der Zuwachs von P^2 bei der Berücksichtigung zusätzlicher Prädiktoren signifikant von null verschieden ist, die zusätzlichen Prädiktoren also einen signifikanten Einfluß haben. Enthält eines der beiden Modelle keine Prädiktoren, dann prüft der Test das Zusammenhangsmaß P^2 aus Gleichung (14).

Die Zahl der Freiheitsgrade der Chiquadratverteilung ergibt sich bei diesem Test durch die Differenz der Parameter in den beiden Regressionsmodellen. Wenn Lo den Wert der Minimierungsfunktion (9) bei dem Modell mit der geringeren Anzahl von Prädiktoren bezeichnet und L1 den Wert der Funktion beim Modell mit der größeren Anzahl von Prädiktoren und ko die Anzahl der Prädiktoren in Modell 0 und k_1 die Anzahl in Modell 1, dann gilt:

$$(16) \quad chi^2 = L_0 - L_1 ; \quad df = k_1 - k_0$$

In unserem Anwendungsbeispiel beträgt die Differenz der Minimierungsfunktionen zwischen dem Nullmodell ohne Prädiktoren und dem Modell mit zwei Prädiktoren 543.35 (= 1188.32 - 644.97). Bei zwei zusätzlichen Regressionskoeffizienten gegenüber dem Nullmodell (df=*2)* ist dieser Wert hochsignifikant. Über die Differenzen der Minimierungsfunktionen sind auch die chi^2 -Werte beim P^2 -Zuwachs in Tabelle 2 ermittelt worden. So beträgt der Wert der Minimierungsfunktion in einer logistischen Regression ohne die VZ-Bewertung, in der die VZ-Notwendigkeit also einziger Regressor ist, 973.93. Die Differenz zum Endmodell beträgt 328.96 (= 973.93 - 644.97). Da ein zusätzlicher Parameter geschätzt wird, beträgt der Freiheitsgrad eins. Statistisch gesehen prüft dieser Test des Netto-Effekts einer einzelnen erklärenden Variable die gleiche Nullhypothese wie der zuerst beschriebene t-Test, nämlich, daß ßk null ist. Tatsächlich sind beide Tests asymptotisch äquivalent.

8 Vgl. MOOD et al. (1974) zu den generellen Eigenschaften von ML-Schätzern und Likelihood-Ratio-Tests.

6. Diskussion

Bei einer dichotomen abhängigen Variable ist die Anwendung des linearen Regressionsmodells oft nicht angemessen. Die Schätzung einer logistischen Regressionsfunktion kann in solchen Fällen eine Alternative sein. Wir haben in diesem Beitrag an einem Beispiel aus der Begleituntersuchung zur Volkszählung demonstriert, wie solche Schätzungen interpretiert werden können.

In unserem Anwendungsbeispiel wird die Boykottabsicht bei der VZ durch zwei Einstellungsvariablen erklärt. Den relativ größeren Einfluß hat dabei eine Variable, die auf einer siebenstufigen Skala die eher affektive Komponente der Einstellung mißt. Von den insgesamt knapp 46% Erklärungskraft des Gesamtmodells werden nicht ganz 28% erst bei Berücksichtigung dieser Variable erreicht. Der zweite Prädiktor mißt eher die instrumentell-kognitive Seite der VZ-Einstellung. Er gibt auf einer dichotomen Skala an, ob ein Befragter die VZ für notwendig erachtet. Ohne diesen Prädiktor sinkt die Erklärungskraft des Modells um gut 2%.

Aus der geringeren relativen Einflußstärke der VZ-Notwendigkeit kann nicht gefolgert werden, daß dieser Prädiktor keinen Einfluß hat. Tatsächlich ist auch der Effekt der VZ-Notwendigkeit hochsignifikant. Es gilt sogar, daß die Wahrscheinlichkeit eines Boykotts grundsätzlich sehr gering ist, wenn die Zählung als notwendig angesehen wird. Etwas überzeichnet ausgedrückt ist es also beinahe eine notwendige aber nicht hinreichende Bedingung für die Boykottabsicht, daß die VZ als nutzloses Unterfangen aufgefaßt wird. Der größere Einfluß der affektiven VZ-Einstellung ergibt sich also vor allem durch die größere Variationsbreite der Antworten auf diese Frage. Dieses Ergebnis weist darauf hin, daß man, wie auch bei einer linearen Regression, die Gesamtheit der Ergebnisse einer Schätzung berücksichtigen sollte und nicht allein formal definierte Effektstärken.

Die Schätzung der Koeffizienten einer binären logistischen Regression lassen sich mit den drei verbreitesten Statistiksystemen BMDP, SAS und SPSSx leicht ermitteln.[9] Im

9 Die Koeffizienten logistischer Regressionen lassen sich natürlich auch mit einigen anderen Programmsystemen wie z.B. GAUSS schätzen.

Anhang wird detailliert die Vorgehensweise bei der Analyse mit SPSSx beschrieben. Die Koeffizientenschätzung mit der Prozedur PLR von BMDP ist im BMDP-Manual ausführlich beschrieben (vgl.: DIXON et al., 1985: 330-344). Ein Beispiel für Programmanweisungen bei einer Analyse mit der Prozedur CATMOD von SAS findet man im SAS-Manual (SAS Inc., 1985: 191) und im Anhang der Arbeit von LONG (1987).[10]

Die breite Verfügbarkeit eines statistischen Modells ist zweifellos ein wichtiges Kriterium für die Anwendung. Wichtiger noch erscheint uns jedoch die methodische und inhaltliche Angemessenheit für eine spezielle Fragestellung. Hier stellt sich insbesondere die Frage nach Alternativen. Grundsätzlich kommen bei einer dichotomen abhängigen Variable andere nichtlineare Regressionsfunktionen, Diskriminanzanalysen oder loglineare Modelle auf Aggregatdatenebene in Frage.

Voraussetzung für eine loglineare Tabellenanalyse sind hinreichend große Fallzahlen in den einzelnen Subgruppen. Die logistische Regression läßt sich in dieser Hinsicht als ein spezielles loglineares Modell auffassen, bei dem unbegrenzt viele Ausprägungen der unabhängigen Variablen möglich sind, so daß die einzelnen Zellen der Tabelle im Extremfall die Häufigkeiten null oder eins haben können. Der 'Trick" der logistischen Regression besteht nun gerade darin, durch die spezifische Form der Regressionskurve sehr starke Restriktionen über die möglichen Parameterwerte eines solchen loglinearen Modells zu verlangen. Durch diese Restriktionen wird zum einen die Anzahl der nichtredundanten Parameter verringert und zum anderen eine Parameterschätzung trotz stellenweise auftretender leerer oder fast leerer Zellenbesetzungen möglich.

Dichotome abhängige Variablen werden oft diskriminanzanalytisch untersucht. Methodologisch gesehen behandeln Diskriminanzanalysen eine etwas andere Fragestellung als Regressionsanalysen. In Regressionsanalysen ist die Form der Regressionsfunktion von Bedeutung. Die Regressionskurve und ihre Parameter werden oft kausal interpretiert als eine Beschreibung des tatsächlichen Wirkungszusammenhangs zwischen abhängiger Variable und deren Beeinflussungsfaktoren. In der Diskriminanzanalyse geht es dagegen darum, einen Fall seiner zugehörigen Klasse korrekt zuzuordnen, d.h. eine optimale Entscheidung bei begrenzten Informationen zu treffen. Bezogen auf unser Anwendungsbeispiel interessiert bei einer Diskriminanzanalyse also nicht, welche konkrete Boykott-

10 Bei der Analyse mit der Prozedur CATMOD ist zu beachten, daß durch die Wahl einer anderen Referenzkategorie alle Regressionskoeffizienten mit einem umgekehrten Vorzeichen geschätzt werden. Bezogen auf unser Beispiel mit der 0/1-kodierten Boykottabsicht schätzt CATMOD nicht die Wahrscheinlichkeit des Boykotts (BOYKOTT = 1) sondern die Wahrscheinlichkeit der Teilnahme (BOYKOTT = 0).

Wahrscheinlichkeit ein Befragter aufweist, sondern ob er letztendlich boykottieren will oder nicht. Diese Fragestellung läßt sich - wie in der klassischen kanonischen Diskriminanzanalyse - auf der Basis eines linearen Modells oder auch eines nichtlinearen Modells behandeln.

Die Betonung der Form der Regressionskurve bei einer Kausalinterpretation führt zu der Frage, ob sich die logistische Funktionsform theoretisch rechtfertigen läßt. Tatsächlich läßt sich in Anlehnung an Arbeiten von McFADDEN (1974) zeigen, daß in unserem Anwendungsbeispiel mit einer logistischen Kurve zu rechnen ist, wenn die geäußerte Boykottabsicht dem Modell der Werterwartungstheorie folgt, die Prädiktoren dann als Nutzenargumente interpretiert werden können und weitere nichtgemessene Nutzenargumente einer Weibullverteilung folgen (vgl. KÜHNEL, 1987).

Auch bei einem Verzicht auf eine explizite Begründung der Form der Regressionskurve sprechen formale methodisch-technische Argumente für die logistische Regressionsfunktion. Zu nennen ist hier zum einen die Flexibilität der Funktionsform (vgl. Abb. 4) und zum anderen die relativ leichte Berechenbarkeit der Schätzwerte für die logistische Regression. Trotzdem kann natürlich nicht - wie auch bei allen anderen Verfahren - prinzipiell die Möglichkeit der Fehlspezifikation des Schätzmodells ausgeschlossen werden. Ein "narrensicheres" Analyseverfahren für alle denkbaren Fragestellungen gibt es nicht und kann es auch nicht geben.

Anhang

Die Berechnung von binären logistischen Regressionen mit SPSSx

Die Schätzung logistischer Regressionsfunktionen kann in SPSSx durch die Prozeduren PROBIT und CNLR erfolgen. (Die Beschreibung der Prozedur CNLR (Conditional Non-Linear Regression) findet man im SPSSx-Handbuch im Kapitel über nichtlineare Regression.) Im Handbuch zur Version 3 von SPSSx (SPSS Inc., 1988) sind für beide Prozeduren Anwendungsbeispiele zur logistischen Regression beschrieben. Da diese jedoch mehr auf die speziellen Eigenschaften der Prozeduren abstellen, möchten wir im folgenden zusammenfassend darstellen, wie man zu den für die Interpretation einer logistischen Regression notwendigen Informationen gelangt.

a) Binäre logistische Regression mit PROBIT

Die Prozedur **PROBIT** ist **primär** für die Analyse von **Aggregatdaten** konzepiert, ermöglicht allerdings auch die logistische Regression von Individualdaten. Aufgrund der Aggregatdatenorientierung wird beim Aufruf der Prozedur erwartet, daß für die Ausprägungskombinationen der unabhängigen Variablen jeweils zwei Informationen verfügbar sind: zum einen die Anzahl der Fälle in der Stichprobe mit einer bestimmten Ausprägungskombination der erklärenden Variablen und zum anderen, bezogen auf diese Fälle, jeweils die Anzahl der Fälle mit der Ausprägung eins bei der 0/1-kodierten abhängigen Variable. Die Syntax der Programmanweisung erwartet daher zwei Variablen, wobei die erste Variable die Anzahl der Fälle mit der Ausprägung 1 auf der abhängigen Variable angibt und die zweite Variable nach dem Schlüsselworrt "OF" die jeweilige Bezugszahl.

Bei einer Analyse auf der Mikroebene der einzelnen Fälle ist die Bezugszahl gerade der einzelne Fall, also grundsätzlich eine Konstante mit dem Wert eins. Die Anzahl der Fälle mit der Ausprägung eins auf der abhängigen Variable kann dann nur 1 oder 0 sein. Sie entspricht gerade dem Wert der abhängigen Variable für den jeweiligen Fall. Um mit PROBIT die Parameter einer logistischen Regression zu schätzen, muß daher im Datensatz die abhängige Variable 0/1-kodiert vorliegen und zusätzlich eine Konstante mit dem Wert eins existieren, was über eine COMPUTE-Anweisung realisiert werden kann.

Nach der Spezifikation der beiden Häufigkeiten folgt auf das Schlüsselwort "WITH" eine Variablenliste mit allen erklärenden Variablen. Nach dieser Liste müssen noch einige für die logistische Regression notwendige Spezifikationen angegeben werden. Wenn also etwa BOYKOTT der Variablenname der 0/1-kodierten abhängigen Variable ist, VZEINST und VZNOTW die Namen zweier erklärender Variablen und CONST der Variablenname einer zu erzeugenden Konstante mit dem Wert eins, dann würde die Spezifikation der Regressionsgleichung von BOYKOTT auf VZEINST und VZNOTW mit folgenden SPSSx-Anweisungen durchgeführt:

```
COMPUTE     CONST = 1
PROBIT      BOYKOTT OF CONST WITH VZEINST VZNOTW
            /MODEL = LOGIT /LOG = NONE /PRINT = NONE
```

Die erklärenden Variablen VZEINST und VZNOTW werden vom Programm als metrische Variablen interpretiert. Im Falle polytom-nominalskalierter Prädiktoren müßten diese zunächst in dichotome Dummy-Variablen aufgelöst werden.

Die nachfolgenden Spezifikationen sind notwendig, da die Voreinstellung der Prozedur PROBIT eine Aggregatdatenanalyse nach dem PROBIT-Modell erwartet. Mit der Spezi-

fikation "/MODEL = LOGIT" wird festgelegt, daß eine logistische Regression berechnet werden soll. Standardmäßig werden alle erklärenden Variablen vor der Analyse mit einer logarithmischen Funktion transformiert. Bei einer logistischen Regression der Individualdaten sind solche Transformationen allerdings nicht unbedingt erwünscht, zumal eine logarithmische Transformation bei Argumenten kleiner oder gleich null nicht definiert ist und zu ungültigen Fällen führen würde. Mit der Spezifikation "/LOG = NONE" ist daher die standardmäßige Transformation auszuschalten. Mit der zusätzlichen Spezifikation "/PRINT = NONE" wird verhindert, daß für jeden Fall der Vorhersagewert (Wert auf der Regressionskurve) und das Residuum ausgedruckt wird, was nur bei Aggregatdatenanalysen sinnvoll bzw. überschaubar ist.

Die Ausgabe der Prozedur gibt zunächst Informationen über die Anzahl der gültigen Fälle und das gewählte Analyseverfahren (Logistisches Modell). Es folgen technische Angaben über die Konvergenz bei der Schätzung der Parameter und schließlich die Werte der Regressionskoeffizienten, Standardfehler und der Quotient aus Koeffizient und Standardfehler (t-Wert). Nach diesen Angaben wird das Ergebniss eines Chiquadrat-Anpassungstests ausgedruckt. Der Test liefert jedoch nur bei Aggregatdatenanalyse und hinreichender Fallzahl in den Aggregatgruppen interpretierbare Ergebnisse. Am Ende der Ausgabe werden noch die Korrelationen der Schätzungen der Regressionsgewichte der Prädiktoren ausgegeben.

Interessant sind für die logistische Regression in erster Linie die Regressionskoeffizienten und die Standardfehler. Dabei ist jedoch zu beachten, daß die ausgedruckten Werte nicht die ursprünglichen Koeffizienten sind, da nicht die Regressionsfunktion (5), sondern wegen der leichteren Vergleichbarkeit und Interpretation von Probit- und Logit-Modellen auf Aggregatdatenebene eine linear transformierte Funktion geschätzt wird. Um die Parameter der ursprünglichen Regressionsfunktion zu bekommen, muß man daher die Ergebnisse des PROBIT-Ausdrucks zurücktransformieren. Dazu müssen alle Regressionskoeffizienten mit 2 multipliziert und von der Regressionskonstante anschließend noch die Zahl 10 abgezogen werden.

Der PROBIT-Ausdruck der Koeffizienten und Standardfehler sieht für unser Anwendungsbeispiel folgendermaßen aus:

Tabelle A1

	Regression Coeff.	Standard Error	Coeff./S.E.
VZEINST	-.71724	.05640	-12.71774
VZNOTW	-.96915	.23901	-4.05483
	Intercept	Standard Error	Intercept/S.E.
	5.78926	.10597	54.63373

Die im Text (Tabelle 2) wiedergegebenen Regressionskoeffizienten ergeben sich dann bei der Schätzung mit PROBIT folgendermaßen:

-1.436 ≈ -.71724 * 2
-1.942 ≈ -.96915 * 2
1.581 ≈ (5.78926 * 2) - 10

Nicht nur die von PROBIT errechneten Koeffizienten, auch die Standardfehler müssen mit 2 multipliziert werden, um die korrekte Zahl zu erhalten. Die Standardfehler haben also die Werte 0.05640 * 2,0.23901 * 2 und 0.10597 * 2. Die t-Werte der Regressionsgewichte bleiben unverändert, weil sowohl im Zähler als auch im Nenner mit 2 multipliziert wird. Bei der Regressionskonstante muß dagegen entweder die rücktransformierte Regressionskonstante durch den rücktransformierten Standardfehler geteilt oder aber vom von PROBIT ausgedruckten t-Wert der Wert 5, geteilt durch den von PROBIT ausgegebenen Standardfehler, abgezogen werden:

7.5 ≈ 1.581 / (0.10597 * 2) = 54.63373 - 5/0.10597

Bei der Berechnung der Regressionskoeffizienten mit den Daten des Anwendungsbeispiels erschien im PROBIT-Ausdruck die Meldung, daß die Schätzung nicht konvergiert hat und das Konvergenzkriterium 0.59547 statt 0.001 beträgt. Tatsächlich stimmen die Schätzwerte jedoch (bis auf unbedeutende Abweichungen) mit der konvergierenden Lösung der Prozedur CNLR (und auch mit der Prozedur CATMOD in SAS) überein, so daß diese Meldung hier keine inhaltliche Bedeutung hat und auf den unterschiedlichen Rechenalgorithmus zurückzuführen ist. Es scheint jedoch ratsam zu sein, bei solchen Meldungen die Ergebnisse mit einer anderen Prozedur zu überprüfen.

Für die Berechnung von P^2, den Zuwachs von P^2 und Likelihood-Ratio-Tests benötigt man den Wert der Minimierungsfunktion (9). Dieser wird jedoch von PROBIT nicht ausgegeben. Falls der Wert nicht mit CNLR berechnet wird, muß man daher den Wert der Minimierungsfunktion in einem zweiten SPSSx-Lauf auf der Basis der geschätzten Koeffizienten mit COMPUTE-Anweisungen berechnen. Dies geschieht am einfachsten, indem man die geschätzten Regressionskoeffizienten in die Gleichung (5b) einsetzt, anschließend Gleichung (5a) berechnet, und dann diese Werte in Gleichung (9) benutzt. Für unser Anwendungsbeispiel ergeben sich so folgende Anweisungen:

```
COMPUTE      U = 1.581 - 1.436 * VZEINST - 1.942 * VZNOTW
COMPUTE      PROB = 1/(1 + Exp(-1 * U))
COMPUTE      L1 = -2 * (BOYKOTT * LN(PROB) + (1-BOYKOTT) * LN(1-PROB))
DESCRIPTIVES VARIABLES = L1 /STATISTICS = SUM
```

Mit der ersten COMPUTE-Anweisung wird der lineare Teil (Gleichung 5b) der logistischen Regressionsfunktion berechnet, mit der zweiten die Linkfunktion (Gleichung 5a) und mit der dritten die Minimierungsfunktion (9). Wenn die abhängige Variable BOYKOTT die Ausprägung 1 hat, wird die Wahrscheinlichkeit des Boykotts (= PROB) in der Minimierungsfunktion berücksichtigt. Wenn die Variable BOYKOTT den Wert 0 hat, wird dagegen die Wahrscheinlichkeit der Teilnahme (= 1-PROB) in der Minimierungsfunktion berücksichtigt. Die nachfolgende SPSSx-Anweisung DESCRIPTIVES berechnet die Summe der Minimierungsfunktion über alle Fälle und gibt das Ergebnis aus.

Auf die gleiche Art lassen sich auch die Werte der Minimierungsfunktion für den P^2-Zuwachs berechnen. Um den Wert für die VZ-Bewertung zu erhalten, berechnet man etwa zunächst eine logistische Regression, in der die VZ-Notwendigkeit einzige erklärende Variable ist. Im zweiten Schritt berechnet man dann mit den COMPUTE-Anweisungen und einer DESCRIPTIVES-Anweisung den Wert der Minimierungsfunktion. Der Wert der Minimierungsfunktion des Modells ohne Prädiktoren, also mit ßo als einzigem Regressionskoeffizienten, kann direkt mit dem Taschenrechner ausgerechnet werden. Wenn n die Anzahl der gültigen Fälle angibt und ni die Anzahl davon, die bei der abhängigen Variable die Ausprägung 0 haben, berechnet sich der Wert der Minimierungsfunktion Lo nach:

$$(17) \quad L_0 = -2 \left(n_1 * \ln (n_1/n) + (n-n_1) * \ln (1-n_1/n) \right)$$

Da in unserem Beispiel 173 von 2061 Befragten boykottieren wollen (Tabelle 1) ergibt sich für Lo der in Tabelle 2 angegebene Wert:

$$L_0 = -2 \left(173 * \ln (173/2061) + (2061-173) * \ln (1-173/2061) \right)$$
$$\approx 1188.32$$

b) binäre logistische Regression mit CNLR

Seit Version 3.0 besteht in SPSSx die Möglichkeit, nahezu beliebige Regressionsfunktionen mit den Prozeduren NLR und CNLR zu berechnen. Für die ML-Schätzung einer logistischen Regression kann allerdings nur die Prozedur CNLR verwendet werden.

Die Spezifikation der logistischen Regression mit CNLR erfolgt in 2 Schritten. Zunächst muß die Regressionsfunktion und die Minimierungsfunktion im sogenannten "MODEL PROGRAM" angegeben werden. Mit der CNLR-Anweisung wird dann die Schätzung der Parameter angefordert. Für unser Anwendungsbeispiel ergeben sich folgende Anweisungen:

```
MODEL       PROGRAM
            B0=0    B1=0    B2=0
COMPUTE     U=B0 + B1 * VZEINST + B2 * VZNOTW
COMPUTE     PROB=1/(1+Exp(-1 * U))
COMPUTE     L1=-2 * (BOYKOTT * LN(PROB) + (1-BOYKOTT) * LN(1-PROB))
CNLR        BOYKOTT WITH VZEINST VZNOTW
            /PRED= PROB /LOSS= L1
```

In der "MODEL PROGRAM"-Anweisung werden die Startwerte der Regressionskoeffizienten spezifiziert. Jeder Regressionskoeffizient wird in einer temporären Variable gespeichert. Da im Beispiel zwei erklärende Variablen verwendet werden, werden eine Regressionskonstante (B0) und zwei Regressionsgewichte (Bl und B2) spezifiziert. Entgegen den Empfehlungen im SPSSx-Handbuch wird allen Startwerten der Wert 0 zugewiesen. Nach unseren Erfahrungen sind diese Startwerte unproblematisch, während man bei Startwerten ungleich null sehr darauf achten muß, daß bei den nachfolgend berechneten Funktionen nicht versehentlich der Bereich der im Computer darstellbaren Zahlen verlassen wird, wenn Wahrscheinlichkeiten von fast null oder fast eins berechnet werden.

In den nächsten drei COMPUTE-Anweisungen wird die Regressionsfunktion über die Gleichungen (5b) und (5a) und darauf aufbauend die Minimierungsfunktion berechnet. Sieht man davon ab, daß statt der Werte der Regressionskoeffizienten die Parameternamen (B0, Bl und B2) eingesetzt werden, sind die drei Anweisungen mit denen identisch, die oben zur Berechnung der Minimierungsfunktion bei einer Schätzung mit PROBIT diskutiert wurden.

Mit der CNLR-Anweisung wird die Regression gestartet. Ähnlich wie bei der PROBIT-Anweisung wird die abhängige Variable (im Beispiel: BOYKOTT) durch das Schlüssel-

wort "WTTH" von den erklärenden Variablen separiert. Durch die Spezifikation "/PRED = PROB" wird dem Programm mitgeteilt, daß die Regressionsfunktion und damit die Vorhersagewerte in der temporären Variable PROB gespeichert werden. Entsprechend besagt die Spezifikation "/LOSS = L1" mit, daß L1 die zu verwendende Minimierungsfunktion ist. Im SPSSx-Jargon wird statt Minimierungsfunktion der Ausdruck "Verlustfunktion" (loss function) verwendet.

Die Druckausgabe besteht im wesentlichen aus einer Tabelle, in der für jeden Minimierungsschritt (jede Iteration) der Wert der Minimierungsfunktion und die Werte der Regressionsparameter wiedergegeben werden. Interessant ist in erster Linie die letzte Zeile dieser Tabelle, die das Endergebnis wiedergibt. Aus den übrigen Tabellenzeilen kann man noch entnehmen, wie der Minimierungsprozeß verläuft. Für unser Anwendungsbeispiel sieht die erste und letzte Zeile der Tabelle so aus:

Tabelle A2

Iteration	Loss funct.	B0	B1	B2
0.1	2857.152678	.000000000	.000000000	.000000000
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
16.1	644.9712204	1.58074531	-1.4358299	-1.9424790

Die Daten der letzten Zeile sind in Tabelle 2 des Hauptteils aufgenommen worden. Insgesamt sind von uns mit CNLR fünf Modelle geschätzt worden:

- das Basismodell, in dem nur die Regressionskonstante geschätzt wird (Modell 0),
- das Modell mit beiden erklärenden Variablen (Modell 1),
- das Modell mit der VZ-Einstellung als einziger erklärender Variable (Modell 2),
- das Modell mit der VZ-Notwendigkeit als einziger erklärender Variable (Modell 3),
- und schließlich ein Modell, in dem zusätzlich zu den beiden erklärenden Variablen ein Interaktionsterm spezifiziert wird (Modell 4).

Die Ergebnisse sind in der folgenden Tabelle zusammengefaßt:

Tabelle A3

Modell	Iter.	Loss funct.	Konstante B0	VZ-Bewertung B1	VZ-Notwendigk. B2	Interaktion B3
0	6.1	1188.32	-2.390	-	-	-
1	16.1	644.97	1.581	-1.436	-1.942	-
2	13.1	670.55	1.561	-1.527	-	-
3	11.1	973.93	-1.617	-	-3.723	-
4	17.1	644.82	1.603	-1.449	-2.248	0.157

Auf der Grundlage der Ergebnisse für die Modelle 0 bis 3 sind für Tabelle 2 des Haupttextes der Wert von P^2, die P^2-Zuwächse sowie die Chiquadrattests berechnet worden. Der Vergleich der Modelle 1 und 4 zeigt ferner, daß ein Interaktionseffekt keine zusätzliche Erklärungskraft bringt.

Mit CNLR ist es nicht möglich, die Standardfehler der Regressionskoeffizienten zu berechnen. Außerdem benötigt die Prozedur relativ viel Rechenzeit, was auch daran liegt, daß die Minimierung der Verlustfunktion numerisch ohne Spezifikation der 1. und 2. Ableitungen erfolgt. Trotzdem sind die Ergebnisse stabil und identisch mit den Schätzungen nach dem mathematisch exakteren Algorithmus in der Prozedur CATMOD in SAS, was auch für einige andere von uns getestete Anwendungsbeispiele zutrifft.

Literatur

Aldrich, J.H. und F.D. Nelson (1984)
Linear Probability, Logit and Probit Models.
Beverly Hills: Sage.

Andress, H.-J. (1986)
GLIM Verallgemeinerte lineare Modelle.
Braunschweig: Vieweg.

Arminger, G. (1983)
Multivariate Analyse von qualitativen abhängigen Variablen mit verallgemeinerten linearen Modellen.
In: Zeitschrift für Soziologie, 12: 49-64.

Dixon, W.J., M.B. Brown, L. Engelman, J.W. Frane, M.A. Hill, R.I. Jenrich und J.D. Toporek (1985)
BMDP Statistical Software.
Berkeley: University of California Press.

Hanushek, E.A. und J.E. Jackson (1977)
Statistical Methods for Social Scientists.
New York: Academic Press.

Hensher, D.A. und L.W. Johnson (1981)
Applied Discrete-Choice Modeling.
New York: Wiley.

Jagodzinski, W. und S.M. Kühnel (1989)
Zur Schätzung der relativen Effekte von Issueorientierungen, Kandidatenpräferenz und langfristiger Parteibindung auf die Wahlabsicht.
Erscheint in: K. Schmidt (Hrsg), Wahlen, Parteieliten, Politische Einstellungen. Neuere Forschungsergebnisse.
Bern: P. Lang (im Druck).

Kühnel, S.M. (1987)
Probleme bei der Modellierung des Entscheidungsverhaltens bei der Volkszählung auf der Basis von Umfragedaten.
Unveröff. Vortragsmanuskript, Köln: Zentralarchiv für empirische Sozialforschung.

Kühnel, S.M. und M. Terwey (1989)
Einflüsse sozialer Konfliktlinien auf das Wahlverhalten im gegenwärtigen Vierparteiensystem der Bundesrepublik.
Erscheint in: Müller, W., P.P. Mohler, B. Erbslöh, M. Wasmer (Hrsg.)
Blickpunkt Gesellschaft. Einstellungen und Verhalten der Bundesbürger.
Opladen: Westdeutscher Verlag (im Druck).

Long, J.S. (1987)
A Graphical Method for the Interpretation of Multinomial Logit Analysis.
In: Sociological Methods and Research: 15 S. 420-466.

McFadden, D. (1974)
Conditional Logit Analysis of Qualitative Choice Behavior.
In: P. Zarembka, Frontiers in Econometrics. New York: Academic Press, S. 105-142.

SAS Institute Inc. (1985),
SAS User's Guide: Statistics. Version 5 Edition.
Cary, NC: SAS Institute Inc.

Scheuch, E.K., L. Graf und S.M. Kühnel, (1988),
Begleituntersuchung zur Volkszählung 1987. Endbericht.
Köln: Zentralarchiv für empirische Sozialforschung.

SPSS Inc. (1988),
SPSS-X User's Guide 3rd Edition.
Chicago: SPSS Inc.